## SEQUENCE LISTING <110> National Food Research Institute <120> Erythrose Reductase, its cDNA and cell which the cDNA express <130> P131238K <160> 7 <210> 1 <211> 1119 <212> DNA <213> Trichosporonoides megachiliensis <400> 1 atg tet tae aaa cag tae ate eee etg aac gae ggt aac aaa ate eet 48 Met Ser Tyr Lys Gln Tyr Ile Pro Leu Asn Asp Gly Asn Lys Ile Pro 5 1 10 15 gee ett gga ttt ggt act tgg caa get gaa eet ggt caa gtg ggt gea 96 Ala Leu Gly Phe Gly Thr Trp Gln Ala Glu Pro Gly Gln Val Gly Ala 20 25 30 agt gtc aag aac gct gtc aag gct ggg tac cgt cat ttg gat ttg gcc 144 Ser Val Lys Asn Ala Val Lys Ala Gly Tyr Arg His Leu Asp Leu Ala 35 40 45 aaa gtg tac caa aac caa tcg gaa att gga gta gca ctt cag gaa ctg 192 Lys Val Tyr Gln Asn Gln Ser Glu Ile Gly Val Ala Leu Gln Glu Leu 50 60 55 ttt gat caa ggt att gtt aaa cgg gaa gat ttg ttt att acg tcc aaa 240 Phe Asp Gln Gly Ile Val Lys Arg Glu Asp Leu Phe Ile Thr Ser Lys 65 70 75 80 gta tgg aat aac cgt cat gct cct gaa cat gtt gag cct gca ttg gac 288

90

95

Val Trp Asn Asn Arg His Ala Pro Glu His Val Glu Pro Ala Leu Asp

85

gaa	aca	ttg	aaa	gaa	ctt	gga	ttg	tcc	tac	ttg	gat	ttg	tac	ttg	att	336
Glu	Thr	Leu	Lys	Glu	Leu	Gly	Leu	Ser	Tyr	Leu	Asp	Leu	Tyr	Leu	Ile	
			100					105					110			
cat	tgg	ccc	gtt	gcg	ttc	aag	ttt	act	acg	cct	caa	gaa	cta	ttc	cct	384
His	Trp	Pro	Val	Ala	Phe	Lys	Phe	Thr	Thr	Pro	Gln	Glu	Leu	Phe	Pro	
		115					120					125				
act	gag	ccg	gat	aac	aag	gaa	ttg	gcc	gcg	att	gat	gat	tca	atc	aag	432
Thr	Glu	Pro	Asp	Asn	Lys	Glu	Leu	Ala	Ala	Ile	Asp	Asp	Ser	Ile	Lys	
	130					135					140					
ttg	gta	gac	act	tgg	aag	gca	gtt	gta	gca	ctc	aaa	aaa	acg	ggt	aag	480
Leu	Val	Asp	Thr	Trp	Lys	Ala	Val	Val	Ala	Leu	Lys	Lys	Thr	Gly	Lys	
145					150					155					160	
acc	aaa	tcc	gtt	ggt	gtg	tcg	aac	ttc	act	acg	gat	ttg	gta	gac	ttg	528
Thr	Lys	Ser	Val	Gly	Val	Ser	Asn	Phe	Thr	Thr	Asp	Leu	Val	Asp	Leu	
				165					170					175		
gtt	gaa	aaa	gcg	tcg	ggg	gaa	cga	ccg	gcg	gtc	aat	cag	att	gaa	gca	576
Val	Glu	Lys	Ala	Ser	Gly	Glu	Arg	Pro	Ala	Val	Asn	Gln	Ile	Glu	Ala	
			180					185					190			
												cac				624
His	Pro	Leu	Leu	Gln	Gln	Asp	Glu	Leu	Val	Ala	His	His	Lys	Ser	Lys	
		195					200					205				
												aat				672
Asn		Val	Ile	Thr	Ala	Tyr	Ser	Pro	Leu	Gly	Asn	Asn	Val	Ser	Gly	
	210					215					220					
												act				720
	Pro	Pro	Leu	Thr		Asn	Pro	Gly	Ile		Ala	Thr	Ala	Lys	Arg	
225					230					235					240	
τta	aat	cat	act	cct	gct	gcg	gtc	ttg	ctt	gca	tgg	ggg	att	caa	cgt	768

Leu Asn His Thr Pro Ala Ala Val Leu Leu Ala Trp Gly Ile Gln Arg	
245 250 255	
gga tac agt gta ttg gtc aag agt gtt aca cet tet ega att gag age	816
Gly Tyr Ser Val Leu Val Lys Ser Val Thr Pro Ser Arg Ile Glu Ser	
260 265 270	
aat tat gat cag att acc ctt tct cct gaa gaa ttc cag aag gtt acg	864
Asn Tyr Asp Gln Ile Thr Leu Ser Pro Glu Glu Phe Gln Lys Val Thr	
275 280 285	
gat ttg atc aag gaa tat ggc gaa agt cgc aac aat att ccg ttg aat	912
Asp Leu Ile Lys Glu Tyr Gly Glu Ser Arg Asn Asn Ile Pro Leu Asn	
290 295 300	
tat aaa cct tca tgg ccc atc agt gtg ttt ggt aca tcg gat gaa gct	960
Tyr Lys Pro Ser Trp Pro Ile Ser Val Phe Gly Thr Ser Asp Glu Ala	
305 310 315 320	
aag get act eat aag att aac ace aac ett tga gtteagtttg ggaactattt	1013
Lys Ala Thr His Lys Ile Asn Thr Asn Leu Stop	
325 330	
$aaagctgctt\ gctggtcaca\ ttattgtcag\ tacctaccat\ gaagaattca\ atattatttt$	1073
acattgtcaa ccattacatg gatccaaaaa aaaaaaaaa aaaaaa 1119	
<210> 2	
<211> 1077	
<212> DNA	
<213> Trichosporonoides megachiliensis	
<400> 2	
atg tee tae aac aag aac ate eet ete aac gae ggt aac tee att eet	48
Met Ser Tyr Asn Lys Asn Ile Pro Leu Asn Asp Gly Asn Ser Ile Pro	

1				5					10					15		
gct	ctt	ggg	tac	ggt	acc	tgg	caa	gca	gaa	cct	ggt	cag	gtc	ggt	gaa	96
Ala	Leu	Gly	Tyr	Gly	Thr	Trp	Gln	Ala	Glu	Pro	Gly	Gln	Val	Gly	Glu	
			20					25					30			
ggt	gtc	aag	$\operatorname{ctc}$	gct	gtt	aag	gct	ggc	tac	cgt	cat	ttg	gac	ttg	gcc	144
Gly	Val	Lys	Leu	Ala	Val	Lys	Ala	Gly	Tyr	Arg	His	Leu	Asp	Leu	Ala	
		35					40					45				
aaa	gta	tac	cag	aac	caa	acc	gag	att	ggc	caa	gct	ctc	aag	gaa	ctg	192
Lys	Val	Tyr	Gln	Asn	Gln	Thr	Glu	Ile	Gly	Gln	Ala	Leu	Lys	Glu	Leu	
	50					55					60					
ttt	gat	gag	ggt	gtt	gtc	aag	$\operatorname{cgt}$	gag	gac	$_{ m ctt}$	ttc	atc	act	tcc	aag	240
Phe	Asp	Glu	Gly	Val	Val	Lys	Arg	Glu	Asp	Leu	Phe	Ile	Thr	Ser	Lys	
65					70					75					80	
ctt	tgg	aac	aac	cgc	cac	gct	$\operatorname{cct}$	gag	cac	gtt	gag	$\operatorname{cct}$	gcg	ctc	gac	288
Leu	Trp	Asn	Asn	Arg	His	Ala	Pro	Glu	His	Val	Glu	${\tt Pro}$	Ala	Leu	Asp	
				85					90					95		
gag	act	ctt	aag	gag	ttg	ggt	cta	tcc	tat	ttg	gac	$\operatorname{ctg}$	tac	ttg	att	336
Glu	Thr	Leu	Lys	Glu	Leu	Gly	Leu	Ser	Tyr	Leu	Asp	Leu	Tyr	Leu	Ile	
			100					105					110			
												gaa				384
His	Trp	Pro	Val	Ala	Phe	Lys	Phe	Thr	Thr	Pro	Asp	Glu	Leu	Leu	Pro	
		115					120					125				
												gat				432
Ala		Pro	Thr	Asn	Lys	Asp	Leu	Ala	Tyr	Ile	Asp	Asp	Ser	Val	Lys	
	130					135					140					
												aag				480
	Ser	Asp	Thr	Trp	Lys	Ala	Val	Val	Ala	Leu	Lys	Lys	Thr	Gly	Lys	
145					150					155					160	

acc	aag	tcg	gtt	ggt	gtt	tcg	aac	ttc	age	act	cgt	ctg	gto	gao	ttg	528
Thr	Lys	Ser	Val	Gly	Val	Ser	Asn	Phe	Ser	Thr	Arg	Leu	Val	Asp	Leu	
				165					170					175		
gtt	gag	gaa	gct	tcg	ggc	gaa	cgt	cct	gcg	gtt	aac	cag	atc	gaa	gct	576
Val	Glu	Glu	Ala	Ser	Gly	Glu	Arg	Pro	Ala	Val	Asn	Gln	Ile	Glu	Ala	
			180					185					190			
cac	ccc	ttg	ttg	caa	caa	gac	gag	ttg	gtt	gct	cac	cat	aag	agc	aag	624
His	Pro	Leu	Leu	Gln	Gln	Asp	Glu	Leu	Val	Ala	His	His	Lys	Ser	Lys	
		195					200					205				
aac	att	gtc	atc	act	gct	tac	agt	ccc	ttg	ggc	aac	aat	gtc	gct	ggt	672
Asn	Ile	Val	Ile	Thr	Ala	Tyr	Ser	Pro	Leu	Gly	Asn	Asn	Val	Ala	Gly	
	210					215					220					
aaa	cca	cct	ctg	act	gag	aat	ccc	ggt	att	gtg	gat	gct	gct	aag	cgt	720
Lys	Pro	Pro	Leu	Thr	Glu	Asn	Pro	Gly	Ile	Val	Asp	Ala	Ala	Lys	Arg	
225					230					235					240	
ctg	aac	cat	act	cct	gct	gct	gtg	$\operatorname{ctc}$	att	gct	tgg	ggt	att	caa	cgc	768
Leu	Asn	His	Thr	Pro	Ala	Ala	Val	Leu	Ile	Ala	Trp	Gly	Ile	Gln	Arg	
				245					250					255		
ggg	tac	agc	gtc	ttg	gtc	aag	tca	gtt	aca	ccc	tca	cgg	att	aag	agt	816
Зlу	Tyr	Ser	Val	Leu	Val	Lys	Ser	Val	Thr	Pro	Ser	Arg	Ile	Lys	Ser	
			260					265					270			
						ctg										864
lsn	Phe	Glu	Gln	Ile	Thr	Leu	Ser	Asp	Glu	Glu	Phe	Gln	Arg	Val	Thr	
		275					280					285				
						ggt										912
sn		Ile	Lys	Glu	Tyr	Gly	Glu	Ser	Arg	Asn	Asn	Val	Pro	Phe	Asn	
	290					295					300					
ac	aag	cct	tcg	tgg	tct	att	gac	gtc	ttt	ggt	acc	cag	tac	gag	gct	960

Туг	· Lys	Pro	Ser	Trp	Ser	· Ile	Ası	Va]	Phe	e Gly	7 Thi	Glr	туг	Glu	ı Ala	
308					310					318					320	
aag	gct	acc	cac	aag	att	aac	gct	taa	tg1	tgcto	tta	taca	aaaa	ıgt		1007
Lys	Ala	Thr	His	Lys	Ιlε	Asn	Ala	. Sto	р							
				325	í											
cgt	ttga	acc	tgta	atgt	gt g	aatg	ttat	c ct	catt	gttg	cat	cato	tca	tcaa	aaaaa	a 1067
aaaaaaaaaa													1077			
<2	.0>	3														
<2	11>	112	1													
<2	L2>	DNA														
<2	13>	Tri	chos	por	ono	ides	me	gach	ili	ens	is					
< 40	00>	3														
atg	tcc	tac	aac	aag	aac	atc	$\operatorname{cct}$	ctc	aac	gac	ggt	aac	tcc	atc	$\operatorname{cct}$	48
Met	Ser	Tyr	Asn	Lys	Asn	Ile	Pro	Leu	Asn	Asp	Gly	Asn	Ser	Ile	Pro	
1				5					10					15		
						tgg										96
Ala	Leu	Gly	Tyr	Gly	Thr	Trp	Gln	Ala	Glu	Pro	Gly	Gln	Val	Gly	Glu	
			20					25					30			
						aaa										144
Gly	Val	Lys	Leu	Ala	Val	Lys	Ala	Gly	Tyr	Arg	His	Leu	Asp	Leu	Ala	
		35					40					45				
						acc										192
Lys		Tyr	Gln	Asn	Gln	Thr	Glu	Ile	Gly	Gln	Ala	Leu	Lys	Glu	Leu	
	50					55					60					
						aag										240
Phe	Asp	Glu	Glv	Va1	Val	Lvs	Arc	Glu	Acn	Lon	Dho	110	The	C	T	

65					70	ı				75	i				80	
ctt	tgg	aac	aac	cgc	cac	gct	cct	gag	cac	gtt	gag	cct	gcg	cto	gac	288
Leu	Trp	Asn	Asn	Arg	His	Ala	Pro	Glu	His	Val	Glu	Pro	Ala	Leu	ı Asp	
				85					90					95		
gag	act	ctt	aag	gag	ctg	ggt	cta	tcc	tat	ctg	gac	ctg	tac	ttg	att	336
Glu	Thr	Leu	Lys	Glu	Leu	Gly	Leu	Ser	Tyr	Leu	Asp	Leu	Tyr	Leu	Ile	
			100					105					110			
cac	tgg	cct	gtt	gct	$\operatorname{ttc}$	aag	ttc	act	act	ccc	gat	gaa	ttg	ctc	cct	384
His	Trp	Pro	Val	Ala	Phe	Lys	Phe	Thr	Thr	Pro	Asp	Glu	Leu	Leu	Pro	
		115					120					125				
gct	gac	cct	acc	aac	aag	gat	ctt	gcc	tac	gtt	gac	gat	tcg	gta	aaa	432
Ala	Asp	Pro	Thr	Asn	Lys	Asp	Leu	Ala	Tyr	Val	Asp	Asp	Ser	Val	Lys	
	130					135					140					
ttg	tec	gac	acc	tgg	aag	gcg	gtc	gtc	gcc	ctg	aag	aag	acg	ggt	aag	480
Leu	Ser	Asp	Thr	Trp	Lys	Ala	Val	Val	Ala	Leu	Lys	Lys	Thr	Gly	Lys	
145					150					155					160	
acc	aag	tcg	gtt	ggt	gtt	tcg	aac	ttc	agc	act	cgt	ctg	gtc	gac	ttg	528
Thr	Lys	Ser	Val	Gly	Val	Ser	Asn	Phe	Ser	Thr	Arg	Leu	Val	Asp	Leu	
				165					170					175		
			gct													576
Val	Glu	Glu	Ala	Ser	Gly	Glu	Arg	Pro	Ala	Val	Asn	Gln	Ile	$\operatorname{Glu}$	Ala	
			180					185					190			
			ttg													624
His	Pro	Leu	Leu	Gln	Gln	Asp	Glu	Leu	Val	Ala	His	His	Lys	Ser	Lys	
		195					200					205				
			atc													672
		Val	Ile	Thr	Ala	Tyr	Ser	Pro	Leu	Gly	Asn	Asn	Val	Ala	Gly	
	210					215					220					

aaa	cca	$\operatorname{cct}$	$\operatorname{ctg}$	act	gag	aac	ccc	ggt	att	gtg	gat	gct	gct	aag	cgt	720
Lys	Pro	${\tt Pro}$	Leu	Thr	Glu	Asn	Pro	Gly	Ile	Val	Asp	Ala	Ala	Lys	Arg	
225					230					235					240	
ttg	aac	$\operatorname{cat}$	$\operatorname{act}$	cct	gct	gct	gtg	$\operatorname{ctc}$	att	gct	tgg	ggt	att	caa	cgc	768
Leu	Asn	His	Thr	Pro	Ala	Ala	Val	Leu	Ile	Ala	Trp	Gly	Ile	Gln	Arg	
				245					250					255		
ggg	tac	agc	gtc	ttg	gtc	aag	tca	gtt	aca	ccc	tca	cgg	atc	aag	agt	816
Gly	Tyr	Ser	Val	Leu	Val	Lys	Ser	Val	Thr	Pro	Ser	Arg	Ile	Lys	Ser	
			260					265					270			
aac	ttt	gaa	cag	atc	act	ctg	tct	gat	gag	gaa	ttc	caa	cgg	gtt	acc	864
Asn	Phe	Glu	Gln	Ile	Thr	Leu	Ser	Asp	Glu	Glu	Phe	Gln	Arg	Val	Thr	
		275					280					285				
aac	ctc	atc	aag	gag	tac	ggt	gag	agc	cgt	aac	aac	gtt	cct	ttc	aat	912
Asn	Leu	Ile	Lys	Glu	Tyr	Gly	Glu	Ser	Arg	Asn	Asn	Val	Pro	Phe	Asn	
	290					295					300					
tac	aag	cct	tcg	tgg	tcc	att	gac	gtc	ttt	ggt	acc	cag	gac	gag	gct	960
Tyr	Lys	Pro	Ser	Trp	Ser	Ile	Asp	Val	Phe	Gly	Thr	Gln	Asp	Glu	Ala	
305					310					315					320	
aag	gct	acc	cac	aag	att	aac	gct	taa	tgtg	ctct	ta t	acaa	aaag	ŗt		1007
Lys	Ala	Thr	His	Lys	Ile	Asn	Ala	Stop	)							
				325												
cgtt	tgaa	icc t	gtaa	tgtg	t ga	atgt	tato	cto	atte	ttg	cato	gtct	ca t	caaa	aaaaa	1067
aaaa	aaaa	iaa a	aaaa	aaaa	a aa	aaaa	aaaa	aaa	aaaa	aaa	aaaa	aaaa	aa a	aaa		1121

<210> 4

<211>

<212> DNA

<213> Artificial Sequence

```
<220>
<223> Designed nucleotide based on erythrose reductase type III
DNA sequence
<220>
<221> unsure
<222> 9
<223> n=a or c or g or t
<400> 4
cargarctnt tygaycaygg 20
<210> 5
<211>
<212> DNA
<213> Artificial Sequence
<220>
<223> Designed nucleotide based on erythrose reductase type III
DNA sequence
<220>
<221> unsure
<222> 3
<223> n=a or c or g or t
<220>
<221> unsure
```

```
<222> 18
<223> n=a or c or g or t
<400> 5
tgngcytcna tytgrttnac 20
<210> 6
<211>
<212> PRT
<213> Trichosporonoides megachiliensis
<400> 6
Gln Glu Leu Phe Asp Gln Gly
 1
<210> 7
<211>
<212> PRT
<213> Trichosporonoides megachiliensis
<400> 7
Val Asn Gln Ile Glu Ala His
 1
                5
```